

Application No.: 09/787,072  
Amendment and Response dated July 13, 2004  
Reply to Office Action of June 30, 2004  
Docket No.: 1368-11PCT/US  
Page 2

**Amendments to the Specification:**

Please delete the Sequence Listing and replace the Sequence Listing with the substitute Sequence Listing enclosed herewith.

Please replace the paragraph beginning at page 3, lines 8, with the following paragraph:

-- The present invention provides a method of identifying a nucleic acid sequence involved in ribosomal frameshifting. The method comprises 1) searching a database of gene sequences to identify sequences which contain the sequence XXX YYY Z, wherein XXX represents GGG, AAA, TTT or CCC, YYY represents AAA or TTT, Z represents A, T, or C and wherein XXXYYYZ is not AAAAAAAA (SEQ ID NO:3) or TTTTTTT (SEQ ID NO:4); and 2) further searching among those sequences identified in step 1 for a sequence encoding a pseudoknot structure which is within eight nucleotides of the sequence identified in step 1. --

Please replace the paragraph beginning at page 4, line 9, with the following paragraph:

-- The present invention provides a method of identifying a nucleic acid sequence involved in ribosomal frameshifting. The method comprises searching a database of gene sequences to identify nucleic acid sequences which contain a slippery site and a pseudoknot structure associated with frameshifting. The method comprises first searching for a slippery site, which is identified by the sequence XXX YYY Z, wherein XXX represents GGG, AAA, TTT or CCC; YYY represents AAA or TTT; Z represents A, T, or C; and wherein XXXYYYZ is not AAAAAAAA (SEQ ID NO:3) or TTTTTTT (SEQ ID NO:4). Further searching is conducted among those sequences containing a slippery site for a sequence encoding a pseudoknot structure which is within eight nucleotides of the slippery site sequence. --

Application No.: 09/787,072  
Amendment and Response dated July 13, 2004  
Reply to Office Action of June 30, 2004  
Docket No.: 1368-11PCT/US  
Page 3

Please replace the paragraph beginning at page 4, line 17, with the following paragraph:

- - The slippery site may have any of the following nucleic acid sequences: GGG AAA A (SEQ ID NO: 5), GGG AAA T (SEQ ID NO:6), GGG AAA C (SEQ ID NO:7), AAA AAA T (SEQ ID NO:8), AAA AAA C (SEQ ID NO:9), TTT AAA A (SEQ ID NO:10), TTT AAA T (SEQ ID NO:11), TTT AAA C (SEQ ID NO:12), CCC AAA A (SEQ ID NO:13), CCC AAA T (SEQ ID NO:14), CCC AAA C (SEQ ID NO:15), GGG TTT A (SEQ ID NO:16), GGG TTT T (SEQ ID NO:17), GGG TTT C (SEQ ID NO:18), AAA TTT A (SEQ ID NO:19), AAA TTT T (SEQ ID NO:20), AAA TTT C (SEQ ID NO:21), TTT TTT A (SEQ ID NO:22), TTT TTT C (SEQ ID NO:23), CCC TTT A (SEQ ID NO:24), CCC TTT T (SEQ ID NO:25) and CCC TTT C (SEQ ID NO:26). --

Please replace the paragraph beginning at page 5, line 16 with the following paragraph:

- - The GenBank *Saccharomyces cerevisiae*, *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Gallus gallus*, *Sus scrofa*, *Drosophila melanogaster*, and *Virus divisions*, and  $2 \times 10^4$  random sequences of  $10^3$  bases (G-C content = 50%) were searched using the following algorithmic structure:

Step 1: Search for XXXYYY (slippery site) where:

XXX = GGG, AAA, TTT or CCC

YYY = AAA or TTT

Z = A, T, or C

AND XXXYYY - AAAAAAA (SEQ ID NO:3) or TTTTTTT  
(SEQ ID NO:4). --

Application No.: 09/787,072  
Amendment and Response dated July 13, 2004  
Reply to Office Action of June 30, 2004  
Docket No.: 1368-11PCT/US  
Page 4

Please replace the paragraph beginning at page 9, line 1 with the following paragraph:

- - The results from the *S. cerevisiae* genome should provide the best estimate of the frequency of motif hits, because 1) it is complete, 2) it is on the same order of magnitude as the random control, 3) it contains the least amount of duplications, and 4) it was sequenced without reading-frame bias. Analysis of this dataset revealed 260 motif hits, approximately 5.2-fold more frequent than random. BLAST analysis revealed that 153 different recognized genes or CDS were represented. Since the yeast genome is estimated to contain approximately 5900 genes, these data suggest that at least 2.55% of the genes in the yeast genome contain at least one consensus programmed -1 ribosomal frameshift signal. Further, since the algorithm limited the size of gap1 and gap2 and disallowed slippery sites of TTTTTTT (SEQ ID NO:4) and AAAAAAA (SEQ ID NO:3), the data probably represent an underestimate of the fraction of motif hits containing yeast genes. - -

Please replace the paragraph beginning at page 9, line 13 with the following paragraph:

- - If a subset of cellular genes utilize programmed -1 ribosomal frameshifting, then specific frameshift signals would be evolutionarily conserved in homologous genes from different organisms. A preliminary comparison of the locations and structures of motif hits in homologous genes in the different databases reveals cases where nearly identical motif hits appear to be conserved. Two such examples, a comparison of Fibrillin 2 in human (SEQ ID NO:27) and mouse (SEQ ID NO:28), and of the Sulfonurea Receptor in humans (SEQ ID NO:29) and rat (SEQ ID NO:30) are shown in Fig. 2. It is notable that whereas the slippery sites and stems of the motifs are highly conserved, the lengths of gap3, which are not expected to play a critical role, are variable in both of these examples. Thus it appears that the biologically important elements of the frameshift signals have been conserved, while the unimportant elements have been allowed to drift. - -